

Temporal Variability and Ignorance in Monte Carlo Contaminant Bioaccumulation Models: A Case Study with Selenium in *Mytilus edulis*

Matthew Spencer,^{1*} Nicholas S. Fisher,² Wen-Xiong Wang,³ and Scott Ferson⁴

Although the parameters for contaminant bioaccumulation models most likely vary over time, lack of data makes it impossible to quantify this variability. As a consequence, Monte Carlo models of contaminant bioaccumulation often treat all parameters as having fixed true values that are unknown. This can lead to biased distributions of predicted contaminant concentrations. This article demonstrates this phenomenon with a case study of selenium accumulation in the mussel *Mytilus edulis* in San Francisco Bay. "Ignorance-only" simulations (in which phytoplankton and bioavailable selenium concentrations are constant over time, but sampled from distributions of field measurements taken at different times), which an analyst might be forced to use due to lack of data, were compared with "variability and ignorance" simulations (sampling phytoplankton and bioavailable selenium concentrations each month). It was found that ignorance-only simulations may underestimate or overestimate the median predicted contaminant concentration at any time, relative to variability and ignorance simulations. However, over a long enough time period (such as the complete seasonal cycle in a seasonal model), treating temporal variability as if it were ignorance at least gave a range of predicted concentrations that enclosed the range predicted by explicit treatment of temporal variability. Comparing the temporal variability in field data with that predicted by simulations may indicate whether the right amount of temporal variability is being included in input variables. Sensitivity analysis combined with biological knowledge suggests which parameters might make important contributions to temporal variability. Temporal variability is potentially more complicated to deal with than other types of stochastic variability, because of the range of time scales over which parameters may vary.

KEY WORDS: Bioaccumulation; temporal variability; ignorance; conservative Monte Carlo simulations

1. INTRODUCTION

Bioaccumulation models require parameters such as ambient contaminant concentrations, feeding rates,

growth rates, and other physiological variables.⁽¹⁻³⁾ Estimates of these parameters are subject to at least two types of uncertainty⁽⁴⁾: lack of measurement and measurement errors (which will be collectively referred to as "ignorance") and temporal fluctuations in true values (which will be referred to as "variability"). Knowledge of a single parameter may be simultaneously affected by both ignorance (because there is only a limited quantity and quality of empirical information) and variability (because the true value of the parameter does not remain constant). There are well-defined techniques for dealing with combinations of ignorance

¹ Department of Biochemistry, University of Cambridge, U.K.

² Marine Sciences Research Center, State University of New York at Stony Brook.

³ Department of Biology, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong.

⁴ Applied Biomathematics, Setauket, NY.

*Address correspondence to Mathew Spencer, Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QW, U.K.; ms379@cam.ac.uk.

and variability.^(4,5) Here, the concern is with the consequences of failing to correctly separate ignorance and variability in Monte Carlo simulations, a situation into which many analysts are forced due to lack of data.

Monte Carlo simulations are increasingly used to deal with uncertainty in bioaccumulation modeling,⁽⁶⁾ yet many studies treat all sources of uncertainty as if they were ignorance. The importance of separating out different types of uncertainty and dealing with them appropriately is well known, although temporal variability can be more complicated to simulate than other types of variability.⁽⁷⁻⁹⁾ The input parameters for bioaccumulation models depend on ecosystem properties such as the abundances and growth rates of species in a food web, but the quantitative and qualitative structure of food webs may show substantial temporal variability.⁽¹⁰⁾ The default assumption should be that temporal variability exists, rather than that it does not exist. If enough data are available, temporal changes in food web structure can be linked to bioaccumulation models to predict changes in contaminant concentrations over time.⁽¹¹⁾ Lack of data may be the main reason why temporal variability is rarely included in bioaccumulation models. In cases in which few data are available, analysts often resort to combining measurements of input variables collected at different times by different workers. For example, consider a set of measurements of phytoplankton density in estuaries. High values may occur (1) because some estuaries always have more phytoplankton than others, (2) because of measurement errors such as differences between laboratories, or (3) because some measurements happen to be taken during phytoplankton blooms. If the last example is treated as ignorance in a Monte Carlo simulation, then the assumption is that phytoplankton biomass might constantly remain as high as the highest values observed during blooms. This is not likely true, but extensive seasonal data sets are needed in order to separate differences between locations, measurement errors, and fluctuations over time. Even if there are not enough data to separate these components of uncertainty, it is helpful to know the consequences that result from treating them all as ignorance.

Does treating variability as ignorance make any difference to the results of a simulation? Identical predictions would not be expected at any single moment, but the overall predicted distribution might be found to be unaffected in the long term. Even if the long-term distribution is affected, concern would not be great if it was known that treating variability as ignorance was always conservative. There are several ways in which

conservatism can be defined, but the most relevant to bioaccumulation problems is to define the more conservative of a pair of methods as the one that gives a higher estimate of the risk of a contaminant concentration being at least as large as some given value. "Conservative" will be used in this sense throughout the article.

To illustrate how different approaches in the treatment of variability may affect the predictions of Monte Carlo bioaccumulation models, a case study of selenium accumulation in the mussel *Mytilus edulis* in San Francisco Bay was used. Data on temporal variability in the concentration of bioavailable selenium, and both seasonal and stochastic temporal variability in the concentration of phytoplankton on which mussels feed were available. Two sets of simulations were conducted. In the ignorance-only simulations, it was assumed that an analyst had exactly the same numerical data for these parameters, but was unaware that individual data points were collected at different times from a single site and, therefore, represented variability (assuming that measurement errors were small). Instead, the analyst treated them as if they represented inaccurate estimates of a single true value, and sampled once at the start of each replicate in a Monte Carlo simulation. In the variability and ignorance simulations, it was assumed that the analyst was aware that these measurements represented variability and treated them accordingly (by sampling at intervals throughout each replicate of a Monte Carlo simulation). No data on temporal variability in other parameters were available, so any uncertainty in these parameters was treated as ignorance (sampling each parameter once at the start of each replicate).

It will be shown that treating variability as ignorance does not give the same distribution of predicted concentrations as that obtained by explicitly including variability. The difference is largest when environmental or physiological parameters show large changes over time and the loss rate of contaminants from the organism is low. In addition, these simulations demonstrate that treating variability as ignorance is not necessarily conservative. Discussion focuses on how comparing the amount of temporal variability predicted by a Monte Carlo simulation with the observed temporal variability in field data can indicate whether uncertainty is being treated appropriately.

2. METHODS

2.1. The Model

The example model consisted of three components: a pool of bioavailable selenium, phytoplank-

ton, and the mussel *M. edulis*. First-order kinetics of selenium exchange between the dissolved phase and mussel tissue were assumed, as well as the fact that mussels also accumulate selenium from ingested phytoplankton, which they ingest at a rate proportional to phytoplankton density. Mussel biomass did not affect phytoplankton density in this model. A bioconcentration factor (BCF) was used to predict selenium accumulation by phytoplankton, and a discrete-time equation was used to describe selenium accumulation in *M. edulis*:

$$S_{t+\Delta t} = \frac{(k_t + h_t g_t P_t B_t) E_t}{l_t + f_t g_t N_t P_t} \{1 - \exp[-(l_t + f_t g_t N_t P_t) \Delta t]\} + S_t \exp[-(l_t + f_t g_t N_t P_t) \Delta t], \quad (1)$$

where S is the concentration of selenium in mussels ($\mu\text{g}\cdot\text{g}^{-1}$), k is the uptake rate of selenium by mussels from the dissolved phase ($\text{L}\cdot\text{g}^{-1}\cdot\text{d}^{-1}$), h is the assimilation efficiency of selenium from phytoplankton ingested by mussels (dimensionless), g is the filtration rate of mussels ($\text{L}\cdot\text{g}^{-1}\cdot\text{d}^{-1}$), P is the concentration of phytoplankton ($\text{g}\cdot\text{L}^{-1}$), B is the bioconcentration factor for selenium by phytoplankton ($\text{L}\cdot\text{g}^{-1}$), E is the bioavailable selenium concentration ($\mu\text{g}\cdot\text{L}^{-1}$), l is the rate at which mussels eliminate selenium (d^{-1}), f is the assimilation efficiency of carbon from phytoplankton ingested by mussels (dimensionless), and N is the net growth efficiency of mussels (dimensionless). Subscript t indicates time in days, and subscript Δt represents the time interval in days. All masses of organisms were measured as dry mass.

2.2. Input Parameters

Wherever possible, parameters were estimated from a single site in San Francisco Bay (San Mateo Bridge, approximate location $37^\circ 36' \text{ N}$, $112^\circ 15' \text{ W}$), where long-term phytoplankton data, ambient selenium concentration, and selenium in mussels have been measured. Values are given in the units stated above, which are not necessarily those used in the original sources. For all parameters except phytoplankton concentration, limited information was available on the shapes of distributions, so uniform distributions were used with the observed ranges. Treating each distribution as having an unknown shape would have been more realistic,⁽¹²⁾ but the assumption of uniform distributions can effectively illustrate our points without introducing unnecessary complexity.

For several parameters relating to phytoplankton (B , f , and h), estimates from diatoms are at the high end of the range of values for all phytoplankton. Results using only the estimates from diatoms, which often dominate the phytoplankton of estuaries, will be briefly considered.

E (bioavailable selenium concentration): Selenite is likely to be the main bioavailable selenium species.⁽¹³⁾ Dissolved selenite concentrations at San Mateo under high and low flow conditions (April and September, respectively) were 0.017 and 0.040 $\mu\text{g}\cdot\text{L}^{-1}$.⁽¹⁴⁾

k (selenium uptake rate from dissolved phase by mussels): Experimental estimates of selenium uptake rate in *M. edulis* ranged from 0.032 to 0.039 $\text{L}\cdot\text{g}^{-1}\cdot\text{d}^{-1}$.⁽¹⁵⁾

l (elimination of selenium by mussels): Experimental estimates of selenium elimination in *M. edulis* ranged from 0.022 to 0.026 d^{-1} .⁽¹⁵⁾

g (filtration rate of water by mussels): The filtration rate of *M. edulis* ranged from 52 to 196 $\text{L}\cdot\text{g}^{-1}\cdot\text{d}^{-1}$,⁽¹⁶⁾ and is relatively independent of food quantity and quality at the abundances of phytoplankton that were simulated. It was assumed that all phytoplankton in filtered water were retained.⁽¹⁷⁾

h (assimilation of selenium from ingested phytoplankton by mussels): For *M. edulis* feeding on seven species of phytoplankton, selenium assimilation efficiencies ranged from 0.14 to 0.72, with no strong central tendency.⁽¹⁸⁾ However, the values for chlorophytes tended to be low, and the values for diatoms tended to be high (for diatoms alone, the range was from 0.55 to 0.72).

f (carbon assimilation efficiency of mussels): For *M. edulis* feeding on the same seven species of phytoplankton as above, carbon assimilation efficiencies ranged from 0.08 to 0.85.⁽¹⁸⁾ For diatoms alone, the range was from 0.54 to 0.85.

N (net growth efficiency of mussels): Experimental estimates of net growth efficiency in *M. edulis* ranged from 0.43 to 0.65.⁽¹⁹⁾

B (bioconcentration factor for selenium in phytoplankton): For two species of marine phytoplankton (a dinoflagellate and a diatom), BCFs were 19 and 523 $\text{L}\cdot\text{g}^{-1}$, respectively.⁽²⁰⁾ Data were also available for a coccolithophore, but this group is unlikely to make up more than a negligible proportion of the phytoplankton in San Francisco Bay.⁽²¹⁾ A conver-

sion factor of 4.75 from volume:volume to dry mass BCFs⁽²²⁾ was used.

P (phytoplankton concentration): Estimated chlorophyll *a* concentrations from 1986 through 1993 were obtained from the U.S. Geological Survey database of water quality in San Francisco Bay⁽²¹⁾ (URL for data download: <http://sfbay.wr.usgs.gov/access/wqdata/index.html>). Individual measurements were averaged within months over the three sites closest to San Mateo Bridge to obtain sufficient data with equal sampling intervals. The top 2 m of the water column were averaged, as collections of mussels for selenium analysis were made intertidally,⁽²³⁾ and the tidal range in this area was about 2 m.⁽²¹⁾ Chlorophyll *a* was converted to dry mass using a chlorophyll *a* to carbon ratio of 0.029 and a dry mass to carbon ratio of 3.33.⁽²⁴⁾ There was no obvious long-term trend in the resulting estimates of dry mass, but there was a strong seasonal cycle, with a fairly predictable spring bloom. This cycle can be closely approximated by a set of 12 lognormal distributions, one for each month of the year (Table I). The cumulative distribution and autocorrelation structure of simulated phytoplankton dry mass closely matched the observed data (Fig. 1).

2.3. Dependencies

Carbon and selenium assimilation efficiencies (*f* and *h*, respectively) may be positively associated

Table I. Estimated Distributions of Phytoplankton Dry Mass at San Mateo Bridge, South San Francisco Bay

Month	Mean	SD	Sample size
January	-8.57	0.18	6
February	-8.26	0.33	7
March	-7.03	0.69	8
April	-6.63	0.50	8
May	-7.86	0.33	8
June	-8.17	0.37	8
July	-8.26	0.20	7
August	-8.21	0.17	6
September	-8.42	0.27	5
October	-8.38	0.38	5
November	-8.91	0.39	5
December	-8.77	0.25	7

Note: The values represent means and standard deviations (*SD*) of natural log-transformed dry mass estimates ($\text{g}\cdot\text{L}^{-1}$) from 1986 through 1993. Distributions were assumed to be lognormal.

across species of phytoplankton (Kendall's $\tau_b = 0.52$, $N = 7^{(18)}$). There may also be a negative association between carbon assimilation efficiency and net growth efficiency (Kendall's $\tau_b = -0.25$, $N = 8^{(19)}$). Although this is a weak association, it was included in the simulations. These data were collected under two different sets of experimental conditions.⁽¹⁹⁾ The perfect negative association within each set of conditions ($\tau_b = -1$ in each case, $N = 3$ and 5) suggests that there is likely to be a real relationship. Nelsen's method⁽²⁵⁾ was used to generate correlated uniform distributions of *f*, *h*, and *N*. The partial correlation between *h* and *N*, holding *f* constant, was assumed to be zero.

All other input parameters were assumed to be independent. A correlation between bioavailable selenium (*E*) and phytoplankton biomass (*P*) is plausible, because flow rates are important in determining selenium concentrations in San Francisco Bay,⁽¹⁴⁾ and are also likely to affect phytoplankton blooms.⁽²¹⁾ In general, "stochastic" temporal variability is a short-

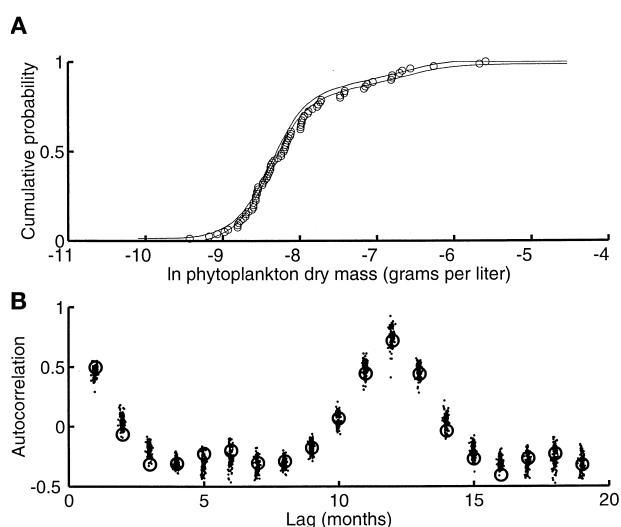


Fig. 1. Observed and simulated phytoplankton distributions at San Mateo, San Francisco Bay. (A) Cumulative probability distributions. Circles are observed data (collected by the U.S. Geological Survey; see Methods section), and the lines are 95% Kolmogorov–Smirnov confidence limits on the simulated distribution. The Kolmogorov–Smirnov confidence interval for the observed distribution is not shown, but it should be much narrower than that given in standard tables, as 24 parameters from this distribution were estimated (12 means and 12 standard deviations) in order to generate the simulated distribution. (B) Autocorrelation coefficients for observed data (circles) and simulations (points, displaced horizontally by small random amounts). All simulations used 100 realizations of 96 months each (the length of the period over which observed data were recorded, although there were 16 missing values in the observed data set).

hand for unknown, underlying physical processes. With a more detailed model, dependencies among temporally varying parameters (such as that between E and P , and many other plausible dependencies caused by physical factors such as temperature and sunlight) would be specified in the model structure. However, with limited knowledge of the system, only empirical correlations can be used. Across sites and sample dates within San Francisco Bay, only a weak correlation was found ($\tau_b = 0.10$, $N = 44$), and, therefore, it was not included in the simulations. However, uptake of selenite during phytoplankton blooms⁽²⁶⁾ might mean that measured ambient concentrations did not accurately affect the amount of selenium in the system.

A positive correlation is also suspected between the assimilation efficiency of selenium from food by mussels (h) and the bioconcentration factor for selenium in phytoplankton (B), as the highest values for both these variables came from diatoms. However, there were too few data to quantify this.

2.4. Observed Selenium Concentrations in Mussels

Predicted selenium concentrations in December were compared with field concentrations at San Mateo between 1986 and 1992, collected by the Mussel Watch Project of the National Status and Trends Program⁽²³⁾ (URL for data download: <http://www-orca.nos.noaa.gov/projects/nsandt/nsandt.html>). One to three composite samples of 30 mussels each were collected intertidally between December and January, homogenized, and analyzed.⁽²³⁾ Ninety-five percent confidence intervals were estimated on the year-to-year coefficient of variation using a nonparametric bootstrap, as follows. A sample (with replacement) containing the observed number of values from the observed set of measurements for each year was taken, and the among-year variance component was estimated using a Model II analysis of variance.⁽²⁷⁾ The year-to-year coefficient of variation was estimated as the square root of the among-year variance component divided by the mean over all years. This process was repeated 10,000 times and the mean, 2.5th, and 97.5th percentiles of the distribution of grand means and year-to-year coefficients of variation were calculated.

2.5. Simulation Methods

A single Monte Carlo simulation was used to generate two types of results from Equation (1), with

identical sets of sampled parameter values. First, phytoplankton biomass (P) and bioavailable selenium concentration (E) were allowed to vary from month to month, but the uncertainty in all other parameters was treated as ignorance. It was suspected that P and E vary over time, because measurements at a number of different times for these variables have been obtained. It was assumed, for simplicity, that measurement error in P and E is much less important than temporal variability, and could be ignored. Of course there may have been temporal variability in other parameters (and the results suggest in which parameters this is likely to be important), but there is not enough information to be able to quantify this. For each of 4,000 replicates, a set of values was sampled for the other parameters, and then Equation (1) was solved over 10 years with a 1-month time step, sampling new values for P and E each month (the variability and ignorance case). The seasonal cycle of months was simulated by sampling from each monthly distribution for phytoplankton biomass in turn (Table I), and using the appropriate number of days for each month as the time interval. The predicted selenium concentration for a given month was measured at the end of the month, so, for example, the January predicted concentration was for the end of January, and used the January distribution of phytoplankton biomass.

Second, we calculated the equilibrium that would occur if the sampled values of P and E for a given month were maintained indefinitely, which corresponded to treating all parameters as ignorance (the ignorance-only case). This is the procedure that would be followed if the temporal variability in P and E was not known. Comparing the results from the variability and ignorance case and the ignorance-only case would, therefore, show the difference between temporal variability and ignorance in P and E . To ensure independence of replicates, month-by-month cumulative distributions of predicted selenium concentrations for each case were constructed, using only the last 12 months of each replicate. The means and year-to-year coefficients of variation for each month in the variability and ignorance case were recorded over the last 7 years of each replicate, to obtain a single mean and year-to-year coefficient of variation for each month from each replicate. These values were recorded over 7 years to match the field data. Preliminary trials showed that cumulative distributions of December predicted concentrations after 12, 24, 36, and 48 months were indistinguishable, indicating approximate stationarity. The time t_p for a given pertur-

bation to decay to a proportion p of its initial value in a model specified by Equation (1) is

$$t_p = -\frac{\ln p}{l + fgNP}. \quad (2)$$

Using the lowest and highest monthly means for P , and the lowest and highest values for other parameters, t_p was about 27 to 207 days for $p = 0.01$. Thus observable effects of initial conditions after 3 years were not expected.

As a sensitivity analysis, the partial Kendall rank correlation coefficients were estimated between the final December predicted selenium concentration from the variability and ignorance case and the sampled value of each input parameter. Parameters whose uncertainty (whether treated as ignorance or variability) would have large effects on predicted selenium concentration would have large absolute partial correlations with the predicted concentration.

The initial selenium concentration in mussels was set at zero, and simulations were started in January. A specially written Monte Carlo routine in Matlab Version 4.2d (The Mathworks, Inc., Natick, MA) was used for all simulations. Equation (1) contains many repeated parameters, but the procedures ensured that each parameter was sampled only once per month (for temporally varying parameters in the variability and ignorance case) or once per replicate (for all other parameters).

3. RESULTS

The bootstrapped mean of annual means from the observed data was $4.87 \mu\text{g}\cdot\text{g}^{-1}$, with a 95% confidence interval from 4.74 to $5.01 \mu\text{g}\cdot\text{g}^{-1}$. Both the ignorance-only simulations ($M = 2.05 \mu\text{g}\cdot\text{g}^{-1}$, from the 2.5th to 97.5th percentile: range = $0.18 \mu\text{g}\cdot\text{g}^{-1}$ – $6.87 \mu\text{g}\cdot\text{g}^{-1}$) and the variability and ignorance simulations ($M = 2.51 \mu\text{g}\cdot\text{g}^{-1}$, from the 2.5th to 97.5th percentile: range = $0.23 \mu\text{g}\cdot\text{g}^{-1}$ – $7.28 \mu\text{g}\cdot\text{g}^{-1}$) generally predicted lower selenium concentrations in mussels than were observed (Fig. 2), although the variability and ignorance simulations were slightly closer to the observed data. It is likely that the model underpredicted the selenium concentration in mussels because it did not take into account the composition of the phytoplankton. Diatoms are often dominant in spring and autumn blooms in estuaries,⁽²¹⁾ and have high selenium bioconcentration factors.⁽²⁰⁾ In addition, the uptake of selenium from diatoms by mussels is relatively efficient.⁽¹⁸⁾ Assuming that all phytoplankton were diatoms gave December predicted

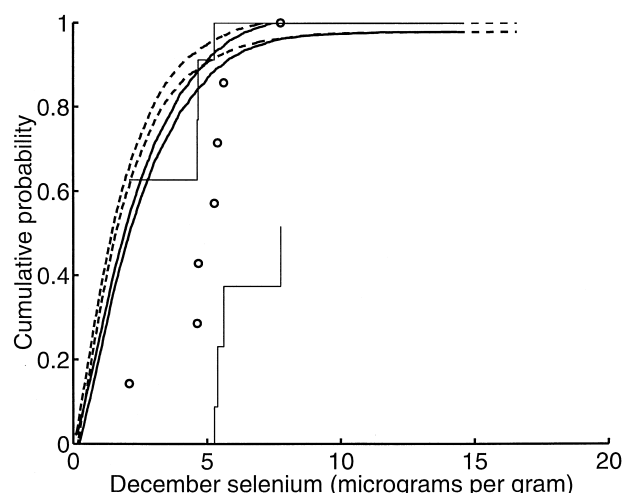
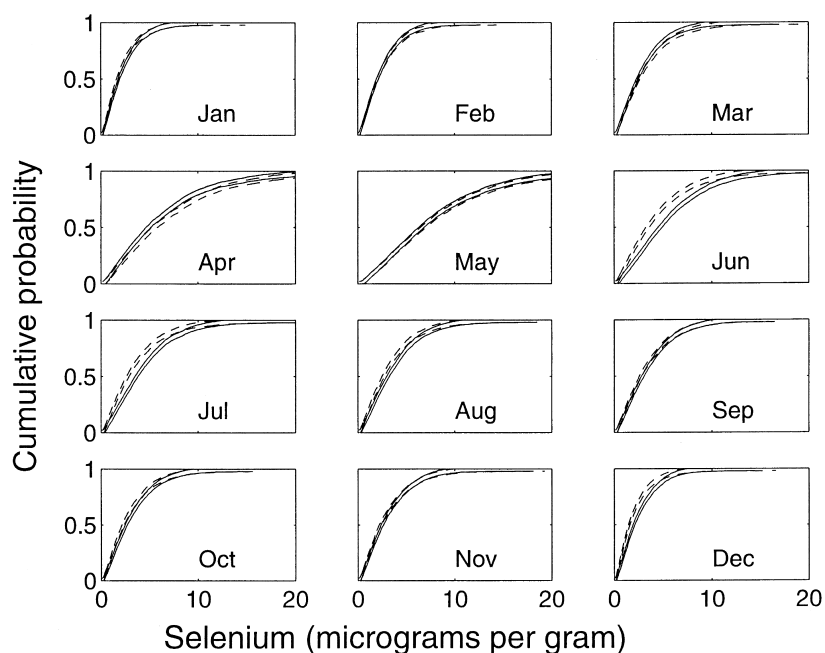


Fig. 2. Observed and predicted cumulative distributions of winter selenium concentration ($\mu\text{g}\cdot\text{g}^{-1}$ dry mass) in *Mytilus edulis* from San Mateo, San Francisco Bay. The circles are the observed data (collected in December and January over 7 years by the National Status and Trends Program; see Methods section) and light solid lines are the 95% Kolmogorov-Smirnov confidence limits. The dashed solid lines are 95% Kolmogorov-Smirnov confidence limits for 4,000 replicates of the December equilibrium solution (the ignorance-only case), and the bold solid lines are 95% Kolmogorov-Smirnov confidence limits for 4,000 replicates of the December solution in which temporal variability in phytoplankton and bioavailable selenium are explicitly included (the variability and ignorance case).

values that were closer to the observed distribution in both cases (ignorance-only: $M = 5.37 \mu\text{g}\cdot\text{g}^{-1}$, range = $1.73 \mu\text{g}\cdot\text{g}^{-1}$ – $11.45 \mu\text{g}\cdot\text{g}^{-1}$; variability and ignorance: $M = 6.40 \mu\text{g}\cdot\text{g}^{-1}$, range = $3.03 \mu\text{g}\cdot\text{g}^{-1}$ – $11.17 \mu\text{g}\cdot\text{g}^{-1}$). This is not a very realistic assumption, but its effect suggests that better information is needed on the relative abundances of different components of the phytoplankton in order to make accurate predictions of selenium bioaccumulation in mussels.

The disagreement between the ignorance-only case and the variability and ignorance case changed with the season (Fig. 3). In four months (June, July, August, and December), the ignorance-only case gave generally lower predicted selenium concentrations in mussels than the variability and ignorance case. In one month (April), the ignorance-only case gave generally higher predicted selenium concentrations than the variability and ignorance case. A simple relationship (Fig. 4) underlies these seasonal changes. When the equilibrium solution increased from the previous month to the current month, the ignorance-only case gave higher predictions than the variability and ignorance case, and vice versa (Fig. 4, top right

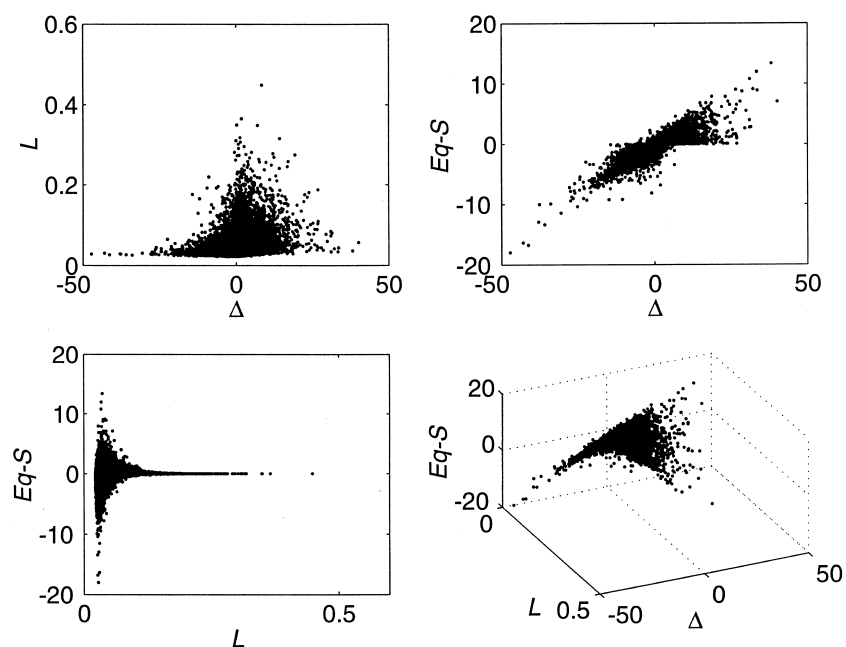
Fig. 3. Cumulative distributions of predicted selenium concentrations ($\mu\text{g}\cdot\text{g}^{-1}$) in mussels for each month of the year, with dashed lines showing the 95% Kolmogorov–Smirnov confidence limits on the ignorance-only case, and solid lines showing the 95% Kolmogorov–Smirnov confidence limits on the variability and ignorance case. Four thousand replicates of each case were run for 10 years each, and predicted values were recorded over the last year. The abscissa has been truncated at $20\ \mu\text{g}\cdot\text{g}^{-1}$ for clarity on all subgraphs (this affects the subgraphs for March, April, May, June, and July; see Table II).



panel), because the ignorance-only case ignored the effect of past selenium concentrations on the current value. Large differences between the ignorance-only case and the variability and ignorance case only occurred when the total loss rate of selenium from mussels [$l + fgNP$ in Equation (1)] was low (Fig. 4, bottom left panel), because a low loss rate implies that the effects of past concentrations persisted for longer

[Equation (2)]. Due to the particular seasonal cycle of phytoplankton biomass occurring in San Francisco Bay (Table I), high loss rates occurred mainly when the change in equilibrium between months was small (Fig. 4, top left panel). The combined relationship between change in equilibrium, total loss rate, and the difference between the ignorance-only and variability and ignorance cases (Fig. 4, bottom right panel)

Fig. 4. Departures of the ignorance-only case from the variability and ignorance case. L is the total loss rate of selenium from mussels at a given time [$l + fgNP$ in Equation (1), d^{-1}], Δ is the equilibrium selenium concentration (predicted by the ignorance-only case) in the current month minus the equilibrium selenium concentration in the previous month ($\mu\text{g}\cdot\text{g}^{-1}$), and $Eq-S$ is the difference between the current equilibrium solution and the current solution from the variability and ignorance case ($\mu\text{g}\cdot\text{g}^{-1}$). The bottom right-hand panel shows the relation among all three variables, and the other panels show the pairwise relations. One point from each of the last 12 months of each of 4,000 replicate simulations over 10 years is plotted.



shows that the equilibrium solution was a reasonable approximation at some, but not all, times.

The ranges of predicted concentrations from the ignorance-only case did not always enclose the ranges of predicted concentrations from the variability and ignorance case (Table II). In June and July, the highest predicted concentrations from the variability and ignorance case were higher than those from the ignorance-only case. This is because the high predicted concentrations in May increased predicted concentrations in subsequent months in the variability and ignorance case. The seasonal cycles in selenium concentration predicted by either model would not be accurate because not enough is known about phytoplankton species composition, but the models demonstrate that explicit treatment of temporal variability does not necessarily give the same results as treating temporal variability as ignorance. Over the entire year (Fig. 5), the range of predicted concentrations from the ignorance-only case ($0.06\text{--}56.88\ \mu\text{g}\cdot\text{g}^{-1}$) was wider than the range of predicted concentrations from the variability and ignorance case ($0.08\text{--}55.60\ \mu\text{g}\cdot\text{g}^{-1}$). The central tendency of the pooled data for the entire year in the ignorance-only case ($M = 3.83\ \mu\text{g}\cdot\text{g}^{-1}$, median = $2.67\ \mu\text{g}\cdot\text{g}^{-1}$) was lower than the central tendency of the pooled data for the variability and ignorance case ($M = 4.01\ \mu\text{g}\cdot\text{g}^{-1}$, median = $2.97\ \mu\text{g}\cdot\text{g}^{-1}$).

The differences in predictions for the entire year between ignorance-only and variability and ignorance simulations were small. However, comparing the observed and predicted coefficients of year-to-year variation in December selenium concentrations

Table II. Ranges by Month for Predicted Selenium Concentrations in Mussels ($\mu\text{g}\cdot\text{g}^{-1}$), for both Ignorance-Only and Variability and Ignorance Simulations^a

Month	Range (ignorance-only)	Range (variability and ignorance)
January	0.06–14.9	0.08–11.59
February	0.09–14.35	0.10–12.69
March	0.06–21.27	0.08–16.30
April	0.14–50.15	0.16–43.05
May	0.15–56.88	0.22–55.60
June	0.10–24.40	0.20–30.25
July	0.07–22.60	0.12–23.81
August	0.09–18.89	0.13–17.99
September	0.09–17.18	0.11–15.98
October	0.09–15.76	0.12–15.18
November	0.10–19.28	0.12–18.10
December	0.06–16.57	0.10–14.35

^a Four thousand replicates in each case.

shows that there was much more temporal variability in the real system than had been included in the model (Fig. 6: predicted mean coefficient of variation = 0.21, 95% confidence interval = 0.10–0.35; observed mean bootstrapped coefficient of variation = 0.34, 95% confidence interval = 0.29–0.38). This was not surprising due to the fact that information about temporal variability was available for only two of the nine input parameters. If such information had been available for the other seven parameters, a larger difference between ignorance-only and variability and ignorance cases would be expected. However, a sensitivity analysis can be used to suggest the parameters in which temporal variability may have large effects on predicted selenium concentration.

The partial correlation coefficients between predicted December selenium concentrations and each input parameter (Table III) show that the parameters whose uncertainty had the largest effects on predicted selenium concentrations were (in descending order of importance) B (the bioconcentration factor for selenium in phytoplankton), h (the assimilation efficiency of selenium from ingested phytoplankton by mussels), and g (the filtration rate). These three parameters all had high uncertainty which was treated as

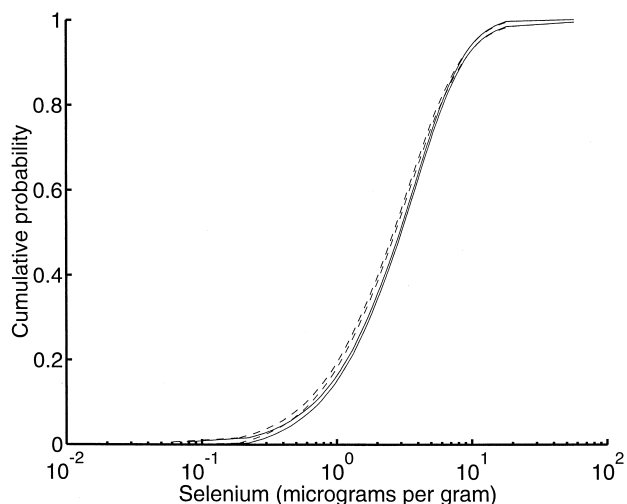


Fig. 5. Cumulative distributions of the natural log of predicted selenium concentrations in mussels pooled over an entire year ($\mu\text{g}\cdot\text{g}^{-1}$), with dashed lines showing the 95% Kolmogorov–Smirnov confidence limits on the ignorance-only case, and solid lines showing the 95% Kolmogorov–Smirnov confidence limits on the variability and ignorance case. Four thousand replicates of each case were run for 10 years each, and predicted values were recorded over the last year (thus each replicate contributed 12 points to the cumulative distribution in each case). Predicted selenium concentrations are plotted on a \log_{10} scale for clarity.

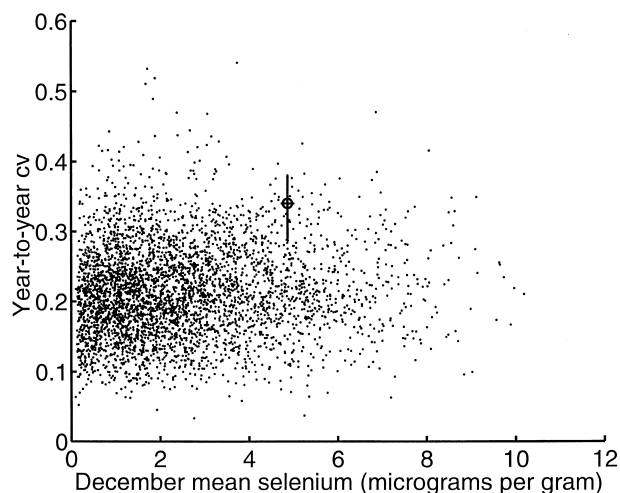


Fig. 6. Observed and predicted year-to-year coefficient of variation (CV) in winter selenium concentration ($\mu\text{g}\cdot\text{g}^{-1}$ dry mass) in *Mytilus edulis* from San Mateo, San Francisco Bay. The circle is from bootstrapping the observed data (7 years, collected by the National Status and Trends Program; see Methods section) and using a Model II analysis of variance to estimate the year-to-year coefficient of variation and grand mean (nonparametric bootstrap, 10,000 replicates), with horizontal and vertical bars giving the 95% confidence interval on the mean and coefficient of variation, respectively. Each dot is the mean and among-year coefficient of variation in predicted December selenium concentration over the last 7 out of 10 years in a Monte Carlo simulation (4,000 replicates).

ignorance, but temporal variability in at least the first two was very likely, because they both depended on the species of phytoplankton. Thus temporal variability in B and h may explain the lower year-to-year coefficient of variation in simulated than observed selenium concentrations. P (the phytoplankton biomass at a given time) and E (the bioavailable selenium concentration), which were treated as varying over time, also had moderately large effects on predicted December selenium concentrations. The uncertainty in the remaining parameters, which was treated as ignorance, was of negligible importance. However, this sensitivity analysis was specific to the amount and treatment of uncertainty, and different conclusions might be drawn if the simulations were repeated after improving parameter estimates.

4. DISCUSSION

Does treating variability as ignorance make any difference to the predictions of Monte Carlo bioaccumulation models? The pooled distribution of predicted selenium concentrations over an entire year differed between the ignorance-only case and the

Table III. Partial Kendall Rank Correlations between Final December Predicted Selenium Concentrations in Mussels and the Sampled Values of Input Parameters^a

Parameter	Partial rank correlation
E	0.1060
k	-0.0002
l	-0.0396
g	0.2841
h	0.3010
f	0.0030
N	-0.0336
B	0.6208
P	0.1378

^aAs defined in Equation (1). See text for further information.

variability and ignorance case (Fig. 5). The appendix shows that in a deterministic seasonal bioaccumulation model, the long-term mean is not generally the same as the mean of the equilibria for each set of seasonal parameter values. We conjecture that the same result holds when there is also stochastic variation. The predicted selenium concentration at any given time and its long-term mean are nonlinear functions of the input parameters, and, in general, the mean of a nonlinear function of random variables is different from the function of the means of those variables.

Is treating variability as ignorance conservative? In June, July, August, and December the predicted values associated with most percentiles were lower in the ignorance-only case than the variability and ignorance case (Fig. 3). Over the whole year (Fig. 5), the median for the ignorance-only case ($2.67 \mu\text{g}\cdot\text{g}^{-1}$) was lower than the median for the variability and ignorance case ($2.97 \mu\text{g}\cdot\text{g}^{-1}$). In this model, treating variability as ignorance was therefore not necessarily more conservative than explicitly including variability. Over the whole year, however, the range of selenium concentrations predicted by the ignorance-only case enclosed the range predicted by the variability-and-ignorance case. The Appendix shows why this occurs in a deterministic seasonal bioaccumulation model. We conjecture that the same result would apply to stochastic models.

When both ignorance and variability are present and can be separated, they can be treated as separate probability spaces,⁽⁵⁾ or as a possibility space for ignorance and a probability space for variability, if there is insufficient information for a probabilistic treatment of ignorance.⁽⁴⁾ In this article, the concern was with the consequences of being unable to make this separation correctly. This may happen when one is unaware that temporal variability exists, or, when one is aware of its existence, but is not able to quantify it.

Our simulations suggest that when it is not possible to separate temporal variability from ignorance and are both treated as ignorance in bioaccumulation models, the maximum predicted contaminant concentration will not be underestimated and the minimum will not be overestimated over a long time scale (longer than the length of any cycle in parameter values). Accurate estimates of any other percentiles, however, may not be obtained. This implies that interval analysis (which does not attempt to make any prediction other than the range of possible values) is more appropriate than Monte Carlo simulation for dealing with a suspected combination of ignorance and variability.⁽⁴⁾ Interval analysis is suboptimal in that it ignores any information available regarding the shapes of distributions. If it is not known, however, whether a distribution results from temporal variability or ignorance, it is not possible to use information about its shape correctly in a Monte Carlo simulation. Because treating variability as ignorance in a Monte Carlo simulation will, in general, affect the central tendency of the output distribution, it can be thought of as a form of structural error. The time scale over which data are needed in order to be sure that the maximum predicted contaminant concentration will not be underestimated and the minimum will not be overestimated depends on the longest cycles in parameter values. Almost 30% of long-term data sets on animal population abundance contain detectable cycles, with many having periods of more than 5 years.⁽²⁸⁾ Thus, ecological risk assessments developed using short-term data may not be reliable in the long term.

There are situations in which treating variability as ignorance will not have major consequences. When the rate at which the effect of past conditions decays [see Equation (2) for the example model] is high, or when the changes in input parameters over a given time interval are sufficiently small, treating variability as ignorance will be a good approximation (Fig. 4), because the system will not be far from equilibrium at any time. One could argue that this study's *Mytilus*-selenium example is such a case, as the differences between the ignorance-only and variability and ignorance simulations were never large. However, the variability and ignorance case probably included substantially less temporal variability than the natural system (Fig. 6), suggesting that most of the uncertainty was still being treated inappropriately, and the potential consequences of variability were still being underestimated. This is hardly surprising, because temporal variability was included in only the two variables for which explicit information was avail-

able. The sensitivity analysis (Table III) suggests that the focus should be on the possibility of temporal variability in bioconcentration factors for selenium in algae, assimilation efficiency of selenium from ingested algae by mussels, and filtration rate. For example, the assumption that all the phytoplankton were diatoms throughout the year made a large difference to predicted selenium concentrations in mussels. In reality, this might be more or less true in spring and autumn, but not in summer.⁽²¹⁾

At first sight, including temporal variability in Monte Carlo simulations of bioaccumulation does not seem computationally difficult; however, the scaling of temporal variability raises a number of problems.⁽⁸⁾ In this study, temporal variability was included on a monthly time scale, as this was the scale on which consistent information on phytoplankton biomass was available. It was assumed that a similar time scale for temporal variability in bioavailable selenium was appropriate. In reality, phytoplankton biomass may show important variation on time scales from hours to decades,⁽²¹⁾ and little information is available on the variability of bioavailable selenium over any time scale. Phytoplankton and bioavailable selenium can be expected to vary on the same scale only if they were governed by the same physical processes. If these physical processes were understood well, attempts could be made to include them in the model structure, rather than adding stochastic variability. Even with little understanding but an abundance of data, it might be possible to establish more suitable time scales on which to sample each parameter in simulations, but a cut-off point would have to be chosen (no shorter than the shortest interval between empirical measurements of the parameter), below which it would be assumed that short-term variability averages out. It would still be necessary to assume that replacing a distribution of time-specific values of a parameter over a given time interval with the distribution of means over the entire time interval would give the same results in a Monte Carlo simulation.⁽⁷⁾ Our results showed that this will not generally be true (also see work by Yool⁽⁹⁾). In practice, confidence in the predictions of Monte Carlo simulations is possible only when there is reasonable certainty that there is no important temporal variability on shorter time scales than that included in the simulation.

ACKNOWLEDGMENTS

The authors are grateful to the Electric Power Research Institute for support to Applied Biomathematics, and to T. P. O'Connor for help in obtaining data on

observed selenium concentrations in mussels. Helen Regan critically read the manuscript, and Resit Akçakaya, Matthew Butcher, Lev Ginzburg, and David Myers commented on earlier versions. Two anonymous referees provided thoughtful and constructive criticism.

APPENDIX: DETERMINISTIC SEASONAL BIOACCUMULATION MODELS

Many simple bioaccumulation models (including the one studied in this article) can be written as a difference equation:

$$C_{t+1} = \frac{\alpha_t}{\beta_t} [1 - \exp(-\beta_t)] + C_t \exp(-\beta_t), \quad (\text{A1})$$

where C_t is the concentration of the contaminant in the organism at time t (dimensions $M_c M_b^{-1}$, where M_c and M_b are the dimensions of contaminant mass and body mass respectively), α_t is the total uptake rate of the contaminant at time t (dimensions $M_c M_b^{-1} T^{-1}$, may include both uptake from the dissolved phase and from ingested particles), and β_t is the total loss rate of contaminant at time t (dimensions T^{-1}). Note that in $\exp(-\beta_t)$, β_t is implicitly multiplied by a constant of 1 time unit. We assume that all $\beta_t > 0$.

Consider a deterministic seasonal form of this model, in which the α_t and β_t are drawn from fixed sequences of length j , repeated indefinitely. Thus $\{\alpha_t, \beta_t\} = \{\alpha_1, \beta_1\}, \{\alpha_2, \beta_2\}, \dots, \{\alpha_j, \beta_j\}$. We will illustrate that for models of this type, (1) the long-term mean may be greater than, less than, or equal to the mean of the equilibria calculated over all pairs of parameters α_t and β_t , for $j \geq 2$; and (2) the range of the set of equilibria calculated over all pairs of parameters α_t and β_t encloses the range of values that occur during a seasonal cycle, for $j \geq 2$.

1. The long-term mean may be greater than, less than, or equal to the mean of the equilibria calculated over all pairs of parameters α_t and β_t , for $j \geq 2$.

Consider a deterministic seasonal model in the form of Equation (A1), where $j = 2$. In the long term, the model converges on a cycle of two values for C :

$$C_1 = \frac{1}{1 - \exp(-\beta_1 - \beta_2)} \times \left(\frac{\alpha_1}{\beta_1} [1 - \exp(-\beta_1)] + \frac{\alpha_2}{\beta_2} [1 - \exp(-\beta_2)] \exp(-\beta_1) \right) \quad (\text{A2})$$

$$C_2 = \frac{1}{1 - \exp(-\beta_1 - \beta_2)} \times \left(\frac{\alpha_2}{\beta_2} [1 - \exp(-\beta_2)] + \frac{\alpha_1}{\beta_1} [1 - \exp(-\beta_1)] \exp(-\beta_2) \right), \quad (\text{A3})$$

so that the long-term mean is $\bar{C}_L = \frac{1}{2}(C_1 + C_2)$. The equilibrium for any pair of parameters α_t and β_t is α_t/β_t (if we assume that the parameters remain constant at these values rather than cycling), so the mean of the equilibria is

$$\bar{C}_E = \frac{1}{2} \left(\frac{\alpha_1}{\beta_1} + \frac{\alpha_2}{\beta_2} \right).$$

The difference between the long-term mean and the mean of the equilibria is then:

$$\bar{C}_E - \bar{C}_L = \frac{1}{2\beta_1} \left(1 - \frac{[1 - \exp(-\beta_1)][1 + \exp(-\beta_2)]}{1 - \exp(-\beta_1 - \beta_2)} \right) + \frac{1}{2\beta_2} \left(1 - \frac{[1 - \exp(-\beta_2)][1 + \exp(-\beta_1)]}{1 - \exp(-\beta_1 - \beta_2)} \right) \quad (\text{A4})$$

This can be either positive, negative, or zero, depending on the parameter values.

2. The range of the set of equilibria calculated over all pairs of parameters α_t and β_t encloses the range of values that occur during a seasonal cycle, for $j \geq 2$.

Consider a model with $j = 2$, in which the equilibrium value for the second pair of parameters is greater than the equilibrium for the first pair of parameters:

$$\frac{\alpha_2}{\beta_2} > \frac{\alpha_1}{\beta_1}.$$

Then, substituting into Equation (A3):

$$C_2 < \frac{1}{1 - \exp(-\beta_1 - \beta_2)} \times \frac{\alpha_2}{\beta_2} [1 - \exp(-\beta_2)] + \frac{\alpha_2}{\beta_2} [1 - \exp(-\beta_1)] \exp(-\beta_2) < \frac{\alpha_2}{\beta_2}, \quad (\text{A5})$$

and similarly,

$$C_1 > \frac{\alpha_1}{\beta_1}.$$

The range of the equilibria is greater than the range of values occurring during the seasonal cycle.

REFERENCES

- Gobas, F. A. P. C. (1993). A model for predicting the bioaccumulation of hydrophobic organic chemicals in aquatic food-webs: Application to Lake Ontario. *Ecological Modelling*, 69, 1–17.
- Thomann, R. V., Connolly, J. P., & Parkerton, T. F. (1992). An equilibrium model of organic chemical accumulation in aquatic

- food webs with sediment interaction. *Environmental Toxicology and Chemistry*, 11, 615–629.
3. Campfens, J., & Mackay, D. (1997). Fugacity-based model of PCB bioaccumulation in complex aquatic food webs. *Environmental Science and Technology*, 31, 577–583.
 4. Ferson, S., & Ginzburg, L. R. (1996). Different methods are needed to propagate ignorance and variability. *Reliability Engineering and Systems Safety*, 54, 133–144.
 5. Helton, J. C. (1997). Uncertainty and sensitivity analysis in the presence of stochastic and subjective uncertainty. *Journal of Statistical Computation and Simulation*, 57, 3–76.
 6. Burkhard, L. P. (1998). Comparison of two models for predicting bioaccumulation of hydrophobic organic chemicals in a Great Lakes food web. *Environmental Toxicology and Chemistry*, 17, 383–393.
 7. Annan, J. D. (1997). On repeated parameter sampling in Monte Carlo simulations. *Ecological Modelling*, 97, 111–115.
 8. Kirchner, T. B. (1998). Time, space, variability, and uncertainty. In M. C. Newman & C. L. Strojan (Eds.), *Risk assessment: Logic and measurement* (pp. 303–323). Chelsea, MI: Ann Arbor Press.
 9. Yool, A. (1999). Comments on the paper: On repeated parameter sampling in Monte Carlo simulations. *Ecological Modelling*, 115, 95–98.
 10. Schoenly, K., & Cohen, J. E. (1991). Temporal variation in food web structure: 16 empirical cases. *Ecological Monographs*, 61, 267–298.
 11. Jackson, L. J. (1996). A simulation model of PCB dynamics in the Lake Ontario pelagic food web. *Ecological Modelling*, 93, 43–56.
 12. Ferson, S., Ginzburg, L., & Akçakaya, R. (manuscript accepted for publication). Whereof one cannot speak: When input distributions are unknown. *Risk Analysis*.
 13. Luoma, S. N., Johns, C., Fisher, N. S., Steinberg, R. S., Oremland, R. S., & Reinfelder, J. R. (1992). Determination of selenium bioavailability to a benthic bivalve from particulate and solute pathways. *Environmental Science and Technology*, 26, 485–492.
 14. Cutter, G. A. (1989). The estuarine behaviour of selenium in San Francisco Bay. *Estuarine, Coastal and Shelf Science*, 28, 13–34.
 15. Wang, W.-X., Fisher, N. S., & Luoma, S. N. (1996). Kinetic determinations of trace element bioaccumulation in the mussel *Mytilus edulis*. *Marine Ecology Progress Series*, 140, 91–113.
 16. Widdows, J., Donkin, P., Brinsley, M. D., Evans, S. V., Salked, P. N., Franklin, A., Law, R. J., & Waldock, M. J. (1995). Scope for growth and contaminant levels in North Sea mussels *Mytilus edulis*. *Marine Ecology Progress Series*, 127, 131–148.
 17. Bayne, B. L., & Newell, R. C. (1983). Physiological energetics of marine molluscs. In A. S. M. Saleuddin & K. M. Wilbur (Eds.), *The molluscs* (pp. 417–515). New York: Academic Press.
 18. Wang, W.-X., & Fisher, N. S. (1996). Assimilation of trace elements and carbon by the mussel *Mytilus edulis*: Effects of food composition. *Limnology and Oceanography*, 41, 197–207.
 19. Kiørboe, T., Møhlenberg, F., & Nøhr, O. (1981). Effect of suspended bottom material on growth and energetics in *Mytilus edulis*. *Marine Biology*, 61, 283–288.
 20. Fisher, N. S., & Reinfelder, J. R. (1995). The trophic transfer of metals in marine systems. In A. Tessier & D. R. Turner (Eds.), *Metal speciation and bioavailability in aquatic systems* (pp. 363–406). Chichester, U.K.: Wiley.
 21. Cloern, J. E. (1996). Phytoplankton bloom dynamics in coastal ecosystems: A review with some general lessons from sustained investigations of San Francisco Bay, California. *Reviews of Geophysics*, 34, 127–168.
 22. Fisher, N. S., Bjerregaard, P., & Fowler, S. W. (1983). Interactions of marine plankton with transuranic elements. 1. Biokinetics of neptunium, plutonium, americium, and californium in phytoplankton. *Limnology and Oceanography*, 28, 432–447.
 23. O'Connor, T. P. (1996). Trends in chemical concentrations in mussels and oysters collected along the U.S. coast from 1986 to 1993. *Marine Environmental Research*, 41, 183–200.
 24. Luoma, S. N., van Geen, A., Lee, B. -G., & Cloern, J. E. (1998). Metal uptake by phytoplankton during a bloom in South San Francisco Bay: Implications for metal cycling in estuaries. *Limnology and Oceanography*, 43, 1007–1016.
 25. Nelsen, R. B. (1986). Properties of a one-parameter family of bivariate distributions with specified marginals. *Communications in Statistics: Theory and Methods*, A15, 3277–3285.
 26. Wrench, J. J., & Measures, C. I. (1982). Temporal variations in dissolved selenium in a coastal ecosystem. *Nature*, 299, 431–433.
 27. Sokal, R. R., & Rohlf, F. J. (1995). *Biometry*. New York: W. H. Freeman.
 28. Kendall, B. E., Prendergast, J., & Bjørnstad, O. N. (1998). The macroecology of population dynamics: Taxonomic and biogeographic patterns in population cycles. *Ecology Letters*, 1, 160–164.